

SEQUENCE LISTING

<110> Suntory Limited

<120> Homeobox gene coding for a protein involved in
differentiation

<130> H773

<160> 8

<210> 1

<211> 1214

<212> DNA

<213> Arabidopsis thaliana

<221> CDS

<222> (36) ... (1010)

<223> Nucleotide sequence coding for a protein involved
in differentiation

<400> 1

ctttagctct cgattatcat cattacacca tcatc atg tcc tcc tca aac aaa	53
Met Ser Ser Ser Asn Lys	
1 5	
aat tgg cca agc atg ttc aaa tcc aaa cct tgc aac aat aat cat cat	101
Asn Trp Pro Ser Met Phe Lys Ser Lys Pro Cys Asn Asn Asn His His	
10 15 20	
cat caa cat gaa atc gat act cca tct tac atg cac tac tct aat tgc	149
His Gln His Glu Ile Asp Thr Pro Ser Tyr Met His Tyr Ser Asn Cys	
25 30 35	
aac cta tca tct tcc ttt tcc tca gat. cgg ata cca gat oct aaa ccg	197
Asn Leu Ser Ser Ser Phe Ser Ser Asp Arg Ile Pro Asp Pro Lys Pro	
40 45 50	
aga tgg aat cct aaa ccg gag cag att agg ata ctc gaa tca atc ttc	245
Arg Trp Asn Pro Lys Pro Glu Gln Ile Arg Ile Leu Glu Ser Ile Phe	
55 60 65 70	

aat tcc ggt act att aac cca cct aga gag gag att caa aga atc cgg	293
Asn Ser Gly Thr Ile Asn Pro Pro Arg Glu Glu Ile Gln Arg Ile Arg	
75 80 85	
atc cgg ctt caa gaa tat ggt caa atc ggt gac gca aac gtg ttt tac	341
Ile Arg Leu Gln Glu Tyr Gly Gln Ile Gly Asp Ala Asn Val Phe Tyr	
90 95 100	
tgg ttt caa aac cgg aaa tct cga gca aaa cac aag ctt cgt gtt cat	389
Trp Phe Gln Asn Arg Lys Ser Arg Ala Lys His Lys Leu Arg Val His	
105 110 115	
cac aaa agc cct aaa atg tca aag aag gac aag acg gtt att cct agt	437
His Lys Ser Pro Lys Met Ser Lys Lys Asp Lys Thr Val Ile Pro Ser	
120 125 130	
act gac gct gat cat tgt ttt ggt ttt gtt aac caa gaa acc gga tta	485
Thr Asp Ala Asp His Cys Phe Gly Phe Val Asn Gln Glu Thr Gly Leu	
135 140 145 150	
tat ccg gtt caa aac aat gag ttg gtg gta acc gaa ccg gcc ggt ttt	533
Tyr Pro Val Gln Asn Asn Glu Leu Val Val Thr Glu Pro Ala Gly Phe	
155 160 165	
cta ttt ccg gtt cat aat gat ccg agc gct gct caa tca gcg ttt ggt	581
Leu Phe Pro Val His Asn Asp Pro Ser Ala Ala Gln Ser Ala Phe Gly	
170 175 180	
ttt ggc gat ttt gtt gta ccg gtg gta acg gaa gaa ggg atg gca ttc	629
Phe Gly Asp Phe Val Val Pro Val Val Thr Glu Glu Gly Met Ala Phe	
185 190 195	
tct acc gtt aat aac ggc gtt aat ttg gag act aac gaa aat ttt gat	677
Ser Thr Val Asn Asn Gly Val Asn Leu Glu Thr Asn Glu Asn Phe Asp	
200 205 210	
aaa att ccg gcg atc aat tta tac ggc gga gat gga aat ggc ggt gga	725
Lys Ile Pro Ala Ile Asn Leu Tyr Gly Gly Asp Gly Asn Gly Gly Gly	
215 220 225 230	
aat tgt ttt cct cct ttg act gtt cca tta acc atc aat caa tct caa	773
Asn Cys Phe Pro Pro Leu Thr Val Pro Leu Thr Ile Asn Gln Ser Gln	
235 240 245	
gaa aaa cga gat gta gga tta tcc ggt ggt gaa gac gtc gga gat aat	821
Glu Lys Arg Asp Val Gly Leu Ser Gly Gly Glu Asp Val Gly Asp Asn	
250 255 260	

gtt tat ccg gtg aga atg acg gtg ttt att aac gag atg cct atc gaa	869
Val Tyr Pro Val Arg Met Thr Val Phe Ile Asn Glu Met Pro Ile Glu	
265 270 275	
gta gtg tct gga tta ttc aac gtt aag gca gct ttc gga aac gat gcc	917
Val Val Ser Gly Leu Phe Asn Val Lys Ala Ala Phe Gly Asn Asp Ala	
280 285 290	
gtt ttg atc aac tcg ttt ggc cag cct att ctt aca gat gaa ttt ggt	965
Val Leu Ile Asn Ser Phe Gly Gln Pro Ile Leu Thr Asp Glu Phe Gly	
295 300 305 310	
gtt act tat caa cct ctc caa aat ggc gca atc tat tat ctt att	1010
Val Thr Tyr Gln Pro Leu Gln Asn Gly Ala Ile Tyr Tyr Leu Ile	
315 320 325	
tagaagatat tgaaaagcaa atgttatggt gctatggata aatattaata taataataaa	1070
agattttctgc gatttattta gttattaatt agataagaat ttcattttctt atctttttaa	1130
tttatgaaca atttacagga catttacatt ttcgagactt tgaaaaataa agaatgaaat	1190
taagttaaaa aaaaaaaaaa aaaa	1214

<210> 2
 <211> 325
 <212> PRT
 <213> Arabidopsis thaliana

<223> Amino acid sequence of protein involved in differentiation

<400> 2

Met Ser Ser Ser Asn Lys Asn Trp Pro Ser Met Phe Lys Ser Lys Pro	
1 5 10 15	
Cys Asn Asn Asn His His His Gln His Glu Ile Asp Thr Pro Ser Tyr	
20 25 30	
Met His Tyr Ser Asn Cys Asn Leu Ser Ser Ser Phe Ser Ser Asp Arg	
35 40 45	
Ile Pro Asp Pro Lys Pro Arg Trp Asn Pro Lys Pro Glu Gln Ile Arg	
50 55 60	
Ile Leu Glu Ser Ile Phe Asn Ser Gly Thr Ile Asn Pro Pro Arg Glu	
65 70 75 80	
Glu Ile Gln Arg Ile Arg Ile Arg Leu Gln Glu Tyr Gly Gln Ile Gly	
85 90 95	

Asp Ala Asn Val Phe Tyr Trp Phe Gln Asn Arg Lys Ser Arg Ala Lys
 100 105 110
 His Lys Leu Arg Val His His Lys Ser Pro Lys Met Ser Lys Lys Asp
 115 120 125
 Lys Thr Val Ile Pro Ser Thr Asp Ala Asp His Cys Phe Gly Phe Val
 130 135 140
 Asn Gln Glu Thr Gly Leu Tyr Pro Val Gln Asn Asn Glu Leu Val Val
 145 150 155 160
 Thr Glu Pro Ala Gly Phe Leu Phe Pro Val His Asn Asp Pro Ser Ala
 165 170 175
 Ala Gln Ser Ala Phe Gly Phe Gly Asp Phe Val Val Pro Val Val Thr
 180 185 190
 Glu Glu Gly Met Ala Phe Ser Thr Val Asn Asn Gly Val Asn Leu Glu
 195 200 205
 Thr Asn Glu Asn Phe Asp Lys Ile Pro Ala Ile Asn Leu Tyr Gly Gly
 210 215 220
 Asp Gly Asn Gly Gly Gly Asn Cys Phe Pro Pro Leu Thr Val Pro Leu
 225 230 235 240
 Thr Ile Asn Gln Ser Gln Glu Lys Arg Asp Val Gly Leu Ser Gly Gly
 245 250 255
 Glu Asp Val Gly Asp Asn Val Tyr Pro Val Arg Met Thr Val Phe Ile
 260 265 270
 Asn Glu Met Pro Ile Glu Val Val Ser Gly Leu Phe Asn Val Lys Ala
 275 280 285
 Ala Phe Gly Asn Asp Ala Val Leu Ile Asn Ser Phe Gly Gln Pro Ile
 290 295 300
 Leu Thr Asp Glu Phe Gly Val Thr Tyr Gln Pro Leu Gln Asn Gly Ala
 305 310 315 320
 Ile Tyr Tyr Leu Ile
 325

<210> 3
 <211> 1518
 <212> DNA
 <213> Arabidopsis thaliana

 <221> CDS
 <222> (152) ... (1285)

<223> Nucleotide sequence coding for a protein involved
in differentiation

<400> 3

```

tttttatttta tcttttccttt agccattctg ttccctgtct cttcctcctc tctttttgac      60
acatcacatc atcatcacat catcattcaa catcaatcat catcatatgc atacacatac      120
atctgtgttc tgcggatcga gttaattagt t atg gct tct tcg aat aga cac      172
                                Met Ala Ser Ser Asn Arg His
                                1             5

tgg cca agc atg ttc aag tcc aaa cct cat ccc cat caa tgg caa cat      220
Trp Pro Ser Met Phe Lys Ser Lys Pro His Pro His Gln Trp Gln His
                10             15             20

gac atc aac tct cct ctc ttg cct tct gct tct cac cga tct tct cct      268
Asp Ile Asn Ser Pro Leu Leu Pro Ser Ala Ser His Arg Ser Ser Pro
                25             30             35

ttc tct tca gga tgt gaa gtg gag agg agt cca gag cca aaa cca aga      316
Phe Ser Ser Gly Cys Glu Val Glu Arg Ser Pro Glu Pro Lys Pro Arg
                40             45             50             55

tgg aat cca aag cca gag cag att cgg ata ctt gaa gca atc ttt aac      364
Trp Asn Pro Lys Pro Glu Gln Ile Arg Ile Leu Glu Ala Ile Phe Asn
                60             65             70

tcc ggg atg gtg aat cct cca aga gag gag atc agg agg att agg gct      412
Ser Gly Met Val Asn Pro Pro Arg Glu Glu Ile Arg Arg Ile Arg Ala
                75             80             85

cag ctt caa gaa tac ggc caa gtc ggt gat gct aac gtc ttc tac tgg      460
Gln Leu Gln Glu Tyr Gly Gln Val Gly Asp Ala Asn Val Phe Tyr Trp
                90             95             100

ttc caa aac cgt aag tcc cgt agt aaa cac aaa ctc cgc ctc ctc cac      508
Phe Gln Asn Arg Lys Ser Arg Ser Lys His Lys Leu Arg Leu Leu His
                105             110             115

aac cac tcc aaa cac tct ctc cct caa acg caa ccg cag ccg cag ccg      556
Asn His Ser Lys His Ser Leu Pro Gln Thr Gln Pro Gln Pro Gln Pro
                120             125             130             135

caa cct tcg gct tcc tct tcc tct tcc tcc tcc tct tcc tcc tcc aaa      604
Gln Pro Ser Ala Ser Ser Ser Ser Ser Ser Ser Ser Ser Ser Ser Lys
                140             145             150

```

tcc acc aaa ccc cga aaa agc aag aac aag aac aac act aat ctc tct	652
Ser Thr Lys Pro Arg Lys Ser Lys Asn Lys Asn Asn Thr Asn Leu Ser	
155 160 165	
ttg ggt ggt agt caa atg atg ggg atg ttt cca ccg gaa ccg gcg ttt	700
Leu Gly Gly Ser Gln Met Met Gly Met Phe Pro Pro Glu Pro Ala Phe	
170 175 180	
ctc ttc ccg gtc tcc act gtc gga ggg ttt gaa ggt atc acc gtc tca	748
Leu Phe Pro Val Ser Thr Val Gly Gly Phe Glu Gly Ile Thr Val Ser	
185 190 195	
tcc caa tta ggg ttt ctc tcc ggt gat atg att gag caa caa aaa ccg	796
Ser Gln Leu Gly Phe Leu Ser Gly Asp Met Ile Glu Gln Gln Lys Pro	
200 205 210 215	
gct cca acg tgt acc gga ctc ctg ctg agt gag atc atg aac ggt agt	844
Ala Pro Thr Cys Thr Gly Leu Leu Leu Ser Glu Ile Met Asn Gly Ser	
220 225 230	
gtg agt tat gga act cat cat caa caa cac ttg agt gag aaa gaa gtt	892
Val Ser Tyr Gly Thr His His Gln Gln His Leu Ser Glu Lys Glu Val	
235 240 245	
gaa gaa atg agg atg aag atg ttg caa cag cca cag act cag att tgt	940
Glu Glu Met Arg Met Lys Met Leu Gln Gln Pro Gln Thr Gln Ile Cys	
250 255 260	
tac gct acc act aat cat caa ata gct tct tac aac aac aac aac aac	988
Tyr Ala Thr Thr Asn His Gln Ile Ala Ser Tyr Asn Asn Asn Asn Asn	
265 270 275	
aac aat aac atc atg ctt cat att cct ccc act act tct act gcc acc	1036
Asn Asn Asn Ile Met Leu His Ile Pro Pro Thr Thr Ser Thr Ala Thr	
280 285 290 295	
act att act act tcg cat tct ctc gct act gtc cca tca act tcg gac	1084
Thr Ile Thr Thr Ser His Ser Leu Ala Thr Val Pro Ser Thr Ser Asp	
300 305 310	
cag ctt caa gtt caa gcg gac gca cga ata aga gtt ttc atc aat gaa	1132
Gln Leu Gln Val Gln Ala Asp Ala Arg Ile Arg Val Phe Ile Asn Glu	
315 320 325	
atg gag ctt gaa gtg agc tca gga ccg ttc aat gtg agg gat gca ttt	1180
Met Glu Leu Glu Val Ser Ser Gly Pro Phe Asn Val Arg Asp Ala Phe	
330 335 340	

ggg gaa gag gtt gtt ctg att aat tcc gcg ggt cag ccc att gtc acc 1228
 Gly Glu Glu Val Val Leu Ile Asn Ser Ala Gly Gln Pro Ile Val Thr
 345 350 355
 gat gaa tat ggc gtc gct ctt cac cct ctt caa cac gga gcc tcg tac 1276
 Asp Glu Tyr Gly Val Ala Leu His Pro Leu Gln His Gly Ala Ser Tyr
 360 365 370 375
 tat ctg atc tagtcgtgtg ggagatttga gtttgaagaa gaaattaaga 1325
 Tyr Leu Ile
 cctgtctctt tctttcacca tctactcgta cgtaggctta aatgttaaga tttataaag 1385
 tattggtttc agttacctgt tgtgacggtg tttatgtatg agtttcggac aacattcaca 1445
 aaactctctc gttaaattgt tgacctataa atatatgatg tgtgtttcat tattaaaaaa 1505
 aaaaaaaaaa aaa 1518

<210> 4
 <211> 378
 <212> PRT
 <213> Arabidopsis thaliana

<223> Amino acid sequence of protein involved in
 differentiation

<400> 4
 Met Ala Ser Ser Asn Arg His Trp Pro Ser Met Phe Lys Ser Lys Pro
 1 5 10 15
 His Pro His Gln Trp Gln His Asp Ile Asn Ser Pro Leu Leu Pro Ser
 20 25 30
 Ala Ser His Arg Ser Ser Pro Phe Ser Ser Gly Cys Glu Val Glu Arg
 35 40 45
 Ser Pro Glu Pro Lys Pro Arg Trp Asn Pro Lys Pro Glu Gln Ile Arg
 50 55 60
 Ile Leu Glu Ala Ile Phe Asn Ser Gly Met Val Asn Pro Pro Arg Glu
 65 70 75 80
 Glu Ile Arg Arg Ile Arg Ala Gln Leu Gln Glu Tyr Gly Gln Val Gly
 85 90 95
 Asp Ala Asn Val Phe Tyr Trp Phe Gln Asn Arg Lys Ser Arg Ser Lys
 100 105 110
 His Lys Leu Arg Leu Leu His Asn His Ser Lys His Ser Leu Pro Gln
 115 120 125

Thr	Gln	Pro	Gln	Pro	Gln	Pro	Gln	Pro	Ser	Ala	Ser	Ser	Ser	Ser	Ser	130	135	140
Ser	Ser	Ser	Ser	Ser	Ser	Lys	Ser	Thr	Lys	Pro	Arg	Lys	Ser	Lys	Asn	145	150	155
Lys	Asn	Asn	Thr	Asn	Leu	Ser	Leu	Gly	Gly	Ser	Gln	Met	Met	Gly	Met	165	170	175
Phe	Pro	Pro	Glu	Pro	Ala	Phe	Leu	Phe	Pro	Val	Ser	Thr	Val	Gly	Gly	180	185	190
Phe	Glu	Gly	Ile	Thr	Val	Ser	Ser	Gln	Leu	Gly	Phe	Leu	Ser	Gly	Asp	195	200	205
Met	Ile	Glu	Gln	Gln	Lys	Pro	Ala	Pro	Thr	Cys	Thr	Gly	Leu	Leu	Leu	210	215	220
Ser	Glu	Ile	Met	Asn	Gly	Ser	Val	Ser	Tyr	Gly	Thr	His	His	Gln	Gln	225	230	235
His	Leu	Ser	Glu	Lys	Glu	Val	Glu	Glu	Met	Arg	Met	Lys	Met	Leu	Gln	245	250	255
Gln	Pro	Gln	Thr	Gln	Ile	Cys	Tyr	Ala	Thr	Thr	Asn	His	Gln	Ile	Ala	260	265	270
Ser	Tyr	Asn	Asn	Asn	Asn	Asn	Asn	Asn	Asn	Ile	Met	Leu	His	Ile	Pro	275	280	285
Pro	Thr	Thr	Ser	Thr	Ala	Thr	Thr	Ile	Thr	Thr	Ser	His	Ser	Leu	Ala	290	295	300
Thr	Val	Pro	Ser	Thr	Ser	Asp	Gln	Leu	Gln	Val	Gln	Ala	Asp	Ala	Arg	305	310	315
Ile	Arg	Val	Phe	Ile	Asn	Glu	Met	Glu	Leu	Glu	Val	Ser	Ser	Gly	Pro	325	330	335
Phe	Asn	Val	Arg	Asp	Ala	Phe	Gly	Glu	Glu	Val	Val	Leu	Ile	Asn	Ser	340	345	350
Ala	Gly	Gln	Pro	Ile	Val	Thr	Asp	Glu	Tyr	Gly	Val	Ala	Leu	His	Pro	355	360	365
Leu	Gln	His	Gly	Ala	Ser	Tyr	Tyr	Leu	Ile							370	375	

<210> 5
 <211> 27
 <212> DNA
 <213> Artificial Sequence

<220>

<221>

<222>

<223> Primer

<400> 5

gaagatctca tcatgtcttc ctcaaac

27

<210> 6

<211> 30

<212> DNA

<213> Artificial Sequence

<220>

<221>

<222>

<223> Primer

<400> 6

cggagctcta aataagataa tagattgcgc

30

<210> 7

<211> 32

<212> DNA

<213> Artificial Sequence

<220>

<221>

<222>

<223> Primer

<400> 7

gctctagaac aatggcttct tcgaatagac ac

32

<210> 8

<211> 32

<212> DNA
<213> Artificial Sequence

<220>
<221>
<222>
<223> Primer

<400> 8
tcccccgggc tgatcagata gtacgaggct cc

32